CS 5522: Artificial Intelligence II

Introduction

Instructor: Wei Xu
Ohio State University

[These slides were adapted from CS188 Intro to AI at UC Berkeley.]
Who am I?

Alan Ritter  
Assistant Professor  
Computer Science and Engineering  
The Ohio State University  
Office: 595 Deere Lab  
Email: ritter.1492@osu.edu  
Google Scholar: https://scholar.google.com/citations?hl=en&user=unX1H3IAAAAAJ  
Semantic Scholar: https://www.semanticscholar.org/author/Alan-Ritter/4655992  

Much of human knowledge is encoded in text. Broadly I am interested in getting computers to better understand natural language. My research uses machine learning to extract knowledge from large heterogeneous corpora, for example, Twitter and the Web. I received my Ph.D. in Computer Science from the University of Washington and was a postdoctoral fellow in the Machine Learning Department at Carnegie Mellon.

Teaching:  
- CSE 5523: Machine Learning and Statistical Pattern Recognition (Fall 2018 - 27972/27973)  
- CSE 5539: Social Media and Text Analytics (Spring 2018)  
- CSE 5522: Survey of Artificial Intelligence II: Advanced Techniques (Fall 2017)  
- CSE 5523: Machine Learning and Statistical Pattern Recognition (Spring 2017)  
- CSE 5525: Speech and Language Processing (Spring 2016)  
- CSE 5539: Web Information Extraction (Fall 2015)  
- CSE 5522: Survey of Artificial Intelligence II: Advanced Techniques (Spring 2015)  
- CSE 5539: NLP and IE for the Social Web (Fall 2014)
Who is the instructor?

http://web.cse.ohio-state.edu/~weixu/

Wei Xu
[phonetic pronunciation: way shoo]
Assistant Professor
Department of Computer Science and Engineering
The Ohio State University
weixu@cse.ohio-state.edu
495 Dreese Lab (2015 Neil Ave, Columbus, OH 43210)

My research lies at the intersections of **machine learning**, **natural language processing**, and **social media**. I focus on designing algorithms for learning semantics from large data for natural language understanding, and generation in particular with stylistic variations. I recently received the NSF CRII Award, Criteo Faculty Research Award, CrowdFlower AI for Everyone Award, Best Paper Award at COLING’18, as well as research funds from DARPA. Previously, I was a postdoctoral researcher at the University of Pennsylvania. I received my PhD in Computer Science from New York University where I was a MacCracken Fellow, MS and BS from Tsinghua University.

I am an area chair for EMNLP 2018 (social media area), COLING 2018 (semantics area), EMNLP 2016 (generation area), a workshop chair for ACL 2017, and the publicity chair for NAACL 2016 and 2018. I also created the [Twitter API tutorial](http://web.cse.ohio-state.edu/~weixu/) and a new course on [Social Media and Text Analytics](http://web.cse.ohio-state.edu/~weixu/).
Where is the instructor?

The 27th International Conference on Computational Linguistics (COLING 2018) will take place in Santa Fe, New Mexico, USA. COLING 2018 will be held at the Santa Fe Community Convention Center from August 20th through 26th 2018.

http://coling2018.org/

Neural Network Models for Paraphrase Identification, Semantic Textual Similarity, Natural Language Inference, and Question Answering

Wuwei Lan and Wei Xu
Department of Computer Science and Engineering
Ohio State University
{lan.105, xu.1265}@osu.edu

Best Paper Award
Today

- What is artificial intelligence?
- What can AI do?
- What is this course?
- Review Some Search Algorithms

- Expectation is that you do **read the textbook**.
Prerequisites

- CSE 3521 or CSE 5521 or CSE 630 or grad standing
- Probability
- Python
- Calculus
- Linear Algebra

- Lots of Math and Programming in Python!
Prerequisites - Probability

- Conditional probability
  \[ P(x|y) = \frac{P(x, y)}{P(y)} \]

- Product rule
  \[ P(x, y) = P(x|y)P(y) \]

- Chain rule
  \[
  P(X_1, X_2, \ldots, X_n) = P(X_1)P(X_2|X_1)P(X_3|X_1, X_2)\ldots
  = \prod_{i=1}^{n} P(X_i|X_1, \ldots, X_{i-1})
  \]

- X, Y independent if and only if: \( \forall x, y : P(x, y) = P(x)P(y) \)

- X and Y are conditionally independent given Z if and only if: \( X \perp Y | Z \)
  \[ \forall x, y, z : P(x, y|z) = P(x|z)P(y|z) \]
Course Website

https://coccoxu.github.io/courses/5522_autum18.html

CSE 5522 Artificial Intelligence II: Advanced Techniques
Advanced concepts, techniques, and applications of artificial intelligence, including knowledge representation, learning, natural language understanding, and vision.

<table>
<thead>
<tr>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wednesday &amp; Friday, 2:20 - 3:40pm</td>
</tr>
<tr>
<td>Place: Caldwell 120</td>
</tr>
<tr>
<td>Instructor: Wei Xu (Office Hours: Wednesday 4:00-5:00pm, Dreese 495)</td>
</tr>
<tr>
<td>TA: TBA</td>
</tr>
</tbody>
</table>

- Example homework, projects, exams, lecture slides are provided;
- subject to change as the autumn 2018 term progresses.
## Course Topics (part 1)

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Required Reading</th>
<th>Suggested Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/22</td>
<td>Course Overview, Reinforcement Learning in NLP Research (Guest Lecture)</td>
<td>3.1-3.4</td>
<td>Microsoft Ms. PacMan</td>
</tr>
<tr>
<td>TBD</td>
<td>Search Review</td>
<td>3.5, 3.6</td>
<td>Pancake Sorting</td>
</tr>
<tr>
<td>TBD</td>
<td>Game Playing 1 - Minimax</td>
<td>5.1, 5.2, 5.3, 5.4</td>
<td>How Checkers was Solved</td>
</tr>
<tr>
<td>TBD</td>
<td>Game Playing 2 - Expectimax and Utilities</td>
<td>5.5, 13.1, 13.2, 16.1, 16.2, 16.3</td>
<td>Rosen’s note on Alpha Beta Pruning</td>
</tr>
</tbody>
</table>

### Reinforcement Learning

<p>| TBD  | Reinforcement Learning 1 - Markov Decision Processes | 17.1, 17.2 |
| TBD  | Reinforcement Learning 2 - Value Iteration and Policy Iteration | 17.3 | Real-life examples of Markov Decision Processes |
| TBD  | Reinforcement Learning 3 - Temporal Difference Learning | 21.1, 21.2, 21.3 | Google AlphaGo |
| TBD  | Reinforcement Learning 4 - Q-Learning | 21.1, 21.2, 21.3 |
| TBD  | Reinforcement Learning 5 - Function Approximation | 21.4 | Sutton &amp; Barto's new reinforcement learning book (Ch. 5, 6, 13) |
| TBD  | Reinforcement Learning 6 - Policy Gradient Methods | 21.5, 21.6 | Deep Q-Learning |</p>
<table>
<thead>
<tr>
<th>Topic</th>
<th>Pages</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoning Under Uncertainty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBD Probability Review</td>
<td>13.3, 13.4, 13.5, 13.6</td>
<td>Andrew Moore’s <a href="#">tutorial</a> and Jean Walrand’s <a href="#">note</a> on Probability</td>
</tr>
<tr>
<td>TBD Markov Models</td>
<td>15.1, 15.2</td>
<td></td>
</tr>
<tr>
<td>TBD Midterm Review, Graphical Models in NLP Research (Guest Lecture)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/10 Midterm (In class - close book and notes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBD Hidden Markov Models 1 - Monitoring and robot localization</td>
<td>15.1, 15.2</td>
<td></td>
</tr>
<tr>
<td>TBD Hidden Markov Models 2 - Particle filtering and resampling</td>
<td>15.3, 15.5</td>
<td></td>
</tr>
<tr>
<td>TBD Bayes Nets 1 - Probabilistic Representations</td>
<td>14.1, 14.2</td>
<td></td>
</tr>
<tr>
<td>TBD Bayes Nets 2 - D-Seperation</td>
<td>14.1, 14.2</td>
<td>Olivier Chapelle’s <a href="#">talk</a> on Bayesian Network Click Model for Web Search</td>
</tr>
<tr>
<td>TBD Bayes Nets 3 - Inference</td>
<td>14.4</td>
<td>David Blei’s <a href="#">talk</a> on Topic Models and User Behavior</td>
</tr>
<tr>
<td>TBD Bayes Nets 4 - Sampling</td>
<td>14.5</td>
<td></td>
</tr>
</tbody>
</table>
## Course Topics (part 3)

<table>
<thead>
<tr>
<th>TBD</th>
<th>Machine Learning and Special Topics</th>
<th>Dependencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD</td>
<td>Perceptron and Naive Bayes</td>
<td>18.6</td>
</tr>
<tr>
<td>TBD</td>
<td>Computer Vision 1 - Applications and CNNs</td>
<td>Facebook Accessibility</td>
</tr>
<tr>
<td>TBD</td>
<td>Computer Vision 2 - Deep Neural Networks and Visualization</td>
<td>Google DeepDream</td>
</tr>
<tr>
<td>TBD</td>
<td>Final Review, Neural Networks in NLP Research (Guest Lecture)</td>
<td>Machine Translation class and book</td>
</tr>
<tr>
<td>TBD</td>
<td>Robotics Research</td>
<td></td>
</tr>
<tr>
<td>TBD</td>
<td>Final Exam (close book and notes)</td>
<td></td>
</tr>
</tbody>
</table>
Grading

Participation (5%)
You will receive credit for engaging in class discussion, asking and answering questions related to the homework on Piazza online discussion board.

Homeworks (15%)
Written homeworks will be very short (one or two exam-style questions) and will be graded in a good/mediocre/incomplete basis. You should be prepared to do regular work each week to keep up with the material and the assignments. Homeworks due before class on day X will include topics we will discuss in class on day X. We will talk about solution in class if people have questions. Homework assignments may NOT be turned in late. Homeworks are NOT accepted by email. There will be 1 grace homework grade per semester, that is, each student receiving full credit for the lowest or a missing homework grade.

Projects (30%)
Programming projects will be in Python, and should be submitted to Carmen by 11:59pm on the day it is due (unless otherwise instructed). Each student will have 3 flexible days to turn in late homework throughout the semester. As an example, you could turn in the first homework 2 days late and the second homework 1 day late without any penalty. After that you will loose 20% for each day the homework is late. Please email your homework to the instructor in case there are any technical issues with submission.

Midterm (20%)
Midterm exam will be close book and notes.

Final Exam (30%)
Final exam will be close book and notes.
Homework Assignments (subject to change as the autumn 2018 term progresses.)

- **Homework 0** - Math and Python Review (due 8/24)
- Homework 1 - Search Review (due TBA)
- **Homework 2** - Markov Decision Processes (due TBA)
- **Homework 3** - Reinforcement Learning: Q-learning (due TBA)
- **Homework 4** - Reinforcement Learning: Feature-based Representations (due TBA)
- **Homework 5** - Probability Review (due TBA)
- **Homework 6** - Hidden Markov Models (due TBA)
- **Homework 7** - Bayes Nets (due TBA)
- **Homework 8** - Naive Bayes (due TBA)

- **Interactive Practice** - Alpha-Beta Pruning
- **Example Exam** - MDPs, Reinforcement Learning, Markov Model, HMM, Bayes Nets
Programming Projects: Berkeley Pac-Man

Pac-Man is a registered trademark of Namco-Bandai Games, used here for educational purposes.

<table>
<thead>
<tr>
<th>Programming Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Project 0 - Unix/Python/Autograder Tutorial (optional)</td>
</tr>
<tr>
<td>• Project 1 - Search (optional)</td>
</tr>
<tr>
<td>• Project 2 - Multi-Agent Search (optional)</td>
</tr>
<tr>
<td>• Project 3 - Reinforcement Learning (due TBA)</td>
</tr>
<tr>
<td>• Project 4 - Ghostbusters (due TBA)</td>
</tr>
<tr>
<td>• Project 5 - Classification (due TBA)</td>
</tr>
</tbody>
</table>
Academic Integrity Policy

Any assignment or exam that you hand in must be your own work (with the exception of group projects). However, talking with others to better understand the material is strongly encouraged. Copying a solution or letting someone copy your solution is cheating. Everything you hand in must be your own words. Code you hand in must be written by you, with the exception of any code provided as part of the assignment. MOSS (Measure of Software Similarity) will be used routinely to detect plagiarism on programming assignments. Any collaboration during an exam is considered cheating. Any student who is caught cheating will be reported to the Committee on Academic Misconduct. Please don’t take a chance - if you are having trouble understanding the material, let us know (asking on Piazza, in class or during office hours), and we will be happy to help.
Resources

- **Piazza** (QA, discussion, and announcements)
- **Carmen** (project submission and restricted resources)
In the last few years, I’ve become increasingly fascinated by artificial intelligence, and in particular our escalating fear of it. It seemed to me that our increasingly holistic relationship with technology and abstract clouds of information was compounding this fear and perhaps edging it
A (Short) History of AI

https://www.youtube.com/watch?v=aygSMgK3BEM
A (Short) History of AI

- **1940-1950: Early days**
  - 1943: McCulloch & Pitts: Boolean circuit model of brain
  - 1950: Turing’s “Computing Machinery and Intelligence”

- **1950–70: Excitement: Look, Ma, no hands!**
  - 1950s: Early AI programs, including Samuel’s checkers program, Newell & Simon’s Logic Theorist, Gelernter’s Geometry Engine
  - 1956: Dartmouth meeting: “Artificial Intelligence” adopted
  - 1965: Robinson’s complete algorithm for logical reasoning

- **1970–90: Knowledge-based approaches**
  - 1969–79: Early development of knowledge-based systems
  - 1980–88: Expert systems industry booms

- **1990–: Statistical approaches**
  - Resurgence of probability, focus on uncertainty
  - General increase in technical depth
  - Agents and learning systems... “AI Spring”?

- **2000–: Where are we now?**
Which of the following can be done at present?

- Play a decent game of table tennis?
- Play a decent game of Jeopardy?
- Drive safely along a curving mountain road?
- Drive safely across campus during the Michigan Game?
- Buy a week’s worth of groceries on the web?
- Buy a week’s worth of groceries at North Market?
- Discover and prove a new mathematical theorem?
- Converse successfully with another person for an hour?
- Perform a surgical operation?
- Put away the dishes and fold the laundry?
- Translate spoken Chinese into spoken English in real time?
- Write an intentionally funny story?

https://www.youtube.com/watch?v=ive4sKkpCqs
One day Joe Bear was hungry. He asked his friend Irving Bird where some honey was. Irving told him there was a beehive in the oak tree. Joe walked to the oak tree. He ate the beehive. The End.

Henry Squirrel was thirsty. He walked over to the river bank where his good friend Bill Bird was sitting. Henry slipped and fell in the river. Gravity drowned. The End.

Once upon a time there was a dishonest fox and a vain crow. One day the crow was sitting in his tree, holding a piece of cheese in his mouth. He noticed that he was holding the piece of cheese. He became hungry, and swallowed the cheese. The fox walked over to the crow. The End.

[Shank, Tale-Spin System, 1984]
Speech and Natural Language Processing

- Speech technologies (e.g. Siri)
  - Automatic speech recognition (ASR)
  - Text-to-speech synthesis (TTS)
  - Dialog systems

- Language processing technologies
  - Question answering
  - Machine translation

- Web search
- Text classification, spam filtering, etc...
- Object and face recognition
- Scene segmentation
- Image classification

Images from Erik Sudderth (left), wikipedia (right)

Demo1: VISION - lec_1_t2_video.flv
Demo2: VISION - lec_1_obj_rec_0.mpg
Robotics

- Robotics
  - Part mech. eng.
  - Part AI
  - Reality much harder than simulations!

- Technologies
  - Vehicles
  - Rescue
  - Soccer!
  - Lots of automation...

- In this class:
  - We ignore mechanical aspects
  - Methods for planning
  - Methods for control

https://www.youtube.com/watch?v=XuPmgeXVKRI
Game Playing

- **Classic Moment: May, '97: Deep Blue vs. Kasparov**
  - First match won against world champion
  - “Intelligent creative” play
  - 200 million board positions per second
  - Humans understood 99.9 of Deep Blue’s moves
  - Can do about the same now with a PC cluster

- **Open question:**
  - How does human cognition deal with the search space explosion of chess?
  - Or: how can humans compete with computers at all??

- **1996: Kasparov Beats Deep Blue**
  “I could feel --- I could smell --- a new kind of intelligence across the table.”

- **1997: Deep Blue Beats Kasparov**
  “Deep Blue hasn't proven anything.”

- Huge game-playing advances recently, e.g. in Go!

Text from Bart Selman, image from IBM’s Deep Blue pages
Decision Making

- Applied AI involves many kinds of automation
  - Scheduling, e.g. airline routing, military
  - Route planning, e.g. Google maps
  - Medical diagnosis
  - Web search engines
  - Spam classifiers
  - Automated help desks
  - Fraud detection
  - Product recommendations
  - ... Lots more!
To Do

- Homework #0
  - Math Review + Python Tutorial
  - Due on in a week
  - Hand in paper copy at the beginning of class next Wednesday

- On the Waiting List:
  - Fill up OSU’s Course Enrollment Permission Form
  - Leave the form at the instructor’s office (slip under the door of DL 495)
  - As space open up, some students will receive an email from the instructor, then pick up the signed form (pinned on the wall outside DL 495)
  - Take the signed form to college office